Wavelet Analysis of Soil Reflectance for the Characterization of Soil Properties

A.V. Bilgili[1], W.D. Hively[2], H. van Es[1], J.Reeves[3], L.Gaston[4]
[1]Crop and Soil Science, Cornell University, NY
[2]USDA-ARS Hydrology and Remote Sensing Laboratory, Beltsville, MD
[3]USDA-ARS Environmental Management and Byproducts Utilization Laboratory, Beltsville, MD
[4]Department of Agronomy and Environmental Management, Louisiana State University, Baton Rouge, LA

ABSTRACT
Wavelet analysis has proven to be effective in many fields including signal processing and digital image analysis. Recently, it has been adapted to spectroscopy, where the reflectance of various materials is measured with respect to wavelength. Reflectance spectra can cover broad wavelength ranges within which wavelength-specific reflectance values can be highly autocorrelated, making the use of traditional statistical procedures impractical for correlating the spectral information with the variable of interest. The spectra also need to processed prior to correlation to remove noise originating from instrument dynamics or atmospheric conditions. Wavelet analysis can provide a good technique to address the aforementioned problems, by reducing the number of necessary wavelengths to the most significant minimum, removing multi-collinearity among the spectral wavelengths, and filtering noise. This project applied wavelet analysis to hyperspectral near-infrared (NIR) and mid-infrared (MIR) reflectance spectra of soil materials, and evaluated its combined use with Multiple Linear Regression Analysis (MLR) and Partial Least Squares Regression (PLSR). Spectral analysis via wavelet decomposition and MLR provided cross validation R^2 values of 0.41, 0.76, 0.48, 0.37, 0.02, 0.83 and 0.50 for soil pH, organic carbon, sand, silt, clay (%), and oxalate extractable Al and Fe (mg kg^{-1}), respectively, with similar results obtained for wavelets + PLSR and PLSR alone. Wavelet analyses successfully reduced the number of wavelengths (predictors) used in the correlation of reflectance spectra with soil properties, and helped with the spectral characterization of certain soil analytes by incorporating different wavelet bases at different scales.

INTRODUCTION
Soil survey and characterization in the agricultural landscape can be hindered by the high cost of chemical and physical laboratory analyses, which are typically time consuming and labor intensive. While reflectance spectroscopy (in the NIR and MIR) has shown promise for rapid and cost-effective characterization of various soil parameters in both laboratory and *in situ* (Viscarra Rossel et al. 2006; Waiser et al. 2007), its accuracy is affected by methods of spectral preprocessing and choice of statistical procedures used to correlate spectra to soil parameters.

Following initial use in geophysical applications in the early 1980s, wavelet transformation and analysis has become commonly employed in other fields, including applications in signal processing (noise removal, improvement of signal/noise ratio, signal compression), remote sensing (feature enhancement and extraction), and digital image analysis (edge detection and land structure classification). It has been reported to outperform other signal processing methods such as Fourier transformation and moving averages because of its ability to localize in both time and frequency domains (Kumar and Georgiou, 1997).

The goals of this project were to: i) enhance characterization of soil properties using wavelet analysis of reflectance spectra; ii) reduce the dimension of necessary predictors (components of reflectance spectra); iii) test the prediction capability of obtained wavelet coefficients in the estimation of soil properties; and iv) compare the estimation accuracy of combined use of wavelet and Multiple Linear Regression (MLR) with results of the Partial Least Square Regression (PLSR) technique commonly used for calibrating soil reflectance to soil properties.

MATERIAL AND METHODS

Soil Samples

Eighty-two soil samples (0-5 cm) were collected from pasture sites in northern Louisiana. The soils included Darley (fine, kaolinitic, thermic Typic Hapludult) and Ruston (fine-loamy, siliceous, thermic Typic Paleudult) series. Samples were air dried, ground to pass a 2 mm screen, and analyzed for pH; organic C (%); soil texture: sand, silt, clay (%); and oxalate-extractable Al and Fe (mg kg^{-1}). The quantity of soil constituents ranged as 4.93 to 6.41 (pH), 0.54 to 6.14 (% C), 45 to 95, 0 to 50, 0 to 15 (texture), 150 to 2708 (Al) and 256 to 28360 (Fe).

VNIR-MID Infrared Spectra

The air-dried ground samples were scanned in the NIR (1000 to 2500 nm) and MIR (2500 to 25000 nm) as described in McCarty et. al. (2002). All spectra were computed as log [1/Reflectance] with 64 co-added scans per spectrum.

Discrete Wavelet Transformation (DWT)

In wavelet transformation, a spectral signal with finite length is decomposed into two series: the first consisting of approximation coefficients (smooth waveform) that capture the overall variation and trend of the spectrum, and the second consisting of wavelet coefficients that capture the fine detail (high frequency part of the signal). The approximation coefficients (A(j,k)) are obtained as an inner product of the signal (f(n)) with a scaled and oriented (shifted) scaling function. The detail coefficients, D(j,k), (wavelet coefficients) are obtained as inner product of signal with scaled and shifted wavelet functions formulated as:

$$A(j,k) = < f(n), \phi_{j,k}(n) > = \sum_{n=0}^{N-1} f(n) x \overline{\phi}_{j,k}(n) \qquad \phi_{j,k}(n) = S_o^{\ j/2} \phi(2^j.n - k)$$

$$D(j,k) = < f(n), \varphi_{j,k}(n) > = \sum_{n=0}^{N-1} f(n) x \overline{\varphi}_{j,k}(n) \qquad \varphi_{j,k}(n) = S_o^{\ j/2} \varphi(2^j.n - k)$$

where N is length of the signal; j is scale and k is shift parameters; and S_o is defined as 2.

Wavelet analysis was performed using Matlab R2006a (The Mathworks Inc. MA,USA). Soil spectra at VNIR range (3124 data points) and MIR range (1868 data points) were decomposed using discrete wavelet transformation with three different wavelet bases, $\varphi_{j,k}$, of Haar, Daubechies 1 (Db1), and Daubechies 6 (Db6). Wavelet coefficients were extracted in this manner (49, 25 and 13 coefficients for NIR and 30, 15 and 8 for MIR spectra at the levels of 6, 7 and 8, respectively) and incorporated into MLR or PLS regression analysis as predictor variables for the estimation of soil parameters. The estimations were validated using repeated one-out cross validation. MLR and PLSR analyses were performed using Unscrambler®V.8.0.5 (CAMO Process AS, Oslo Norway).

RESULTS

The results of PLSR and MLR following wavelet transformation, as well as PLSR without wavelet transformation, for both NIR and MIR spectra, were compared for their estimation accuracy of soil properties using the Root Mean Square Error of Prediction (RMSEP) and the coefficient of determination (R^2) (Table 1).

Table 1. Statistics associated with estimation accuracy of PLSR (original spectra) and discrete wavelet transformation combined with PLSR and MLR.

	VNIR					MID-IR				
	PLSR†	Haar‡	Db1‡	Db6‡	Haar-MLR§	PLSR	Haar	Db1	Db6	Haar-MLR
					RMSEP¶					
OC	0.52	0.51	0.51	0.57	0.66	0.35	0.40	0.39	0.38	0.52
pH	0.21	0.22	0.22	0.22	0.45	0.16	0.18	0.17	0.19	0.23
Sand	7.5	7.4	7.4	6.7	12.3	7.2	7.8	8.0	8.0	9.7
Silt	7.5	7.3	7.3	6.5	12.0	7.1	7.5	7.5	7.7	10.1
Clay	3.3	3.0	3.3	3.3	6.2	3.2	3.2	3.2	3.1	4.4
Al	288	279	279	322	506	255	249	267	263	317
Fe	3545	3822	3826	3791	6898	2980	3222	3212	3062	4855
					R^2					
OC	0.69	0.69	0.68	0.59	0.59	0.84	0.79	0.80	0.81	0.76
pH	0.24	0.13	0.13	0.14	0.05	0.53	0.45	0.54	0.44	0.41
Sand	0.58	0.62	0.62	0.67	0.31	0.60	0.54	0.52	0.52	0.48
Silt	0.49	0.56	0.56	0.63	0.23	0.53	0.50	0.50	0.49	0.37
Clay	0.03	0.02	0.02	0.02	0.05	0.06	0.12	0.11	0.12	0.02
Al	0.83	0.84	0.84	0.81	0.61	0.86	0.88	0.86	0.86	0.83
Fe	0.60	0.55	0.55	0.56	0.45	0.74	0.67	0.68	0.71	0.50

† Partial Least Square Regression with original spectra; ‡Partial Least Squares Regression using wavelet coefficients; ¶ Root Mean Square Error of Prediction; § Multiple Linear Regression using wavelet coefficients; the results for Haar did not differ greatly from results with the other bases, so only Haar is presented.

DISCUSSION AND CONCLUSION

As seen in Table 1, the estimation results using the various methods are similar, and the results from PLSR with original spectra and wavelet coefficients (Haar, Db1 and Db6) follow similar patterns for the various analytes. Applying wavelet transformation before PLSR provided slightly better results in only a few cases. For all statistical methods, MIR spectra provided better results than NIR spectra. The best predictions were derived for Al, Carbon, Fe and Sand.

REFERENCES

Kumar, P and Georgiou, E. 1997. Wavelet Analysis for Geophysical Applications. Reviews of Geophysics, 35(4):385-412.

McCarty, G.W., Reeves, J.B. III, Reeves, V.B., Follett, R.F., and Kimble, J.M. 2002. Mid-infrared and near-infrared diffuse reflectance spectroscopy for measurement of carbon in soils. Soil Sci. Soc. Am. J. 66:640-646.

Viscarra Rossel, R.,Walvoort,D.J.J., McBratney, A.B.,Janik,L.J., Skjemstad,J.O. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. Geoderma 131:59-75.

Waiser, T.H., Morgan, C.L.S., Brown,D.J., Hallmark,C.T. In Situ Characterization of Soil Clay Content with Visible Near-Infrared Diffuse Refl ectance Spectroscopy. SSSAJ,71(2):389-396.

Contact Information: Ali Volkan Bilgili, Crop & Soil Science, Cornell University 1015 Bradfield Hall, Ithaca, NY 14850, USA, Phone: 607-255-1706; Email: avb22@cornell.edu